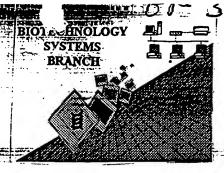# RAW SEQUENCE LISTING
ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: *09/737, 476*

Source: *OIPE*

Date Processed by STIC: *01-08-01*

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY
FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)
PATENTIN 3.0 e-mail help: patin30help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE <u>CHECKER VERSION 3.0 PROGRAM</u>, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0

The Checker Version 3.0 application is a state-of the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.
Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO). Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address:
http://www.uspto.gov/web/offices/pac/checker

Raw Sequence Listing Error Summary

ERROR DETECTED **SUGGESTED CORRECTION** SERIAL NUMBER: _09/737,476_

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1 _____ Wrapped Nucleics
The number/text at the end of each line "wrapped" down to the next line.
This may occur if your file was retrieved in a word processor after creating it.
Please adjust your right margin to .3, as this will prevent "wrapping".

2 _____ Wrapped Aminos
The amino acid number/text at the end of each line "wrapped " down to the next line.
This may occur if your file was retrieved in a word processor after creating it.
Please adjust your right margin to .3, as this will prevent "wrapping".

3 _____ Incorrect Line Length
The rules require that a line not exceed 72 characters in length. This includes spaces.

4 _____ Misaligned Amino Acid Numbering
The numbering under each 5th amino acid is misaligned. This may be caused by the use of tabs between the numbering. It is recommended to delete any tabs and use spacing between the numbers.

5 _____ Non-ASCII
This file was not saved in ASCII (DOS) text, as required by the Sequence Rules.
Please ensure your subsequent submission is saved in ASCII text so that it can be processed.

6 _____ Variable Length
Sequence(s) _____ contain n's or Xaa's which represented more than one residue.
As per the rules, each n or Xaa can only represent a single residue.
Please present the maximum number of each residue having variable length and indicate in the (ix) feature section that some may be missing.

7 _____ PatentIn ver. 2.0 "bug"
A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequence(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies primarily to the mandatory <220>-<223> sections for Artificial or Unknown sequences.

8 _____ Skipped Sequences (OLD RULES)
Sequence(s) _____ missing. If intentional, please use the following format for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X:
(i) SEQUENCE CHARACTERISTICS:(Do not insert any headings under "SEQUENCE CHARACTERISTICS")
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X:
This sequence is intentionally skipped

Please also adjust the "(iii) NUMBER OF SEQUENCES:" response to include the skipped sequence(s).

9 _____ Skipped Sequences (NEW RULES)
Sequence(s) _____ missing. If intentional, please use the following format for each skipped sequence.
<210> sequence id number
<400> sequence id number
000

10 _____ Use of n's or Xaa's (NEW RULES)
Use of n's and/or Xaa's have been detected in the Sequence Listing.
Use of <220> to <223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

11 _____ Use of <213>Organism (NEW RULES)
Sequence(s) _____ are missing this mandatory field or its response.

12 ✓ Use of <220>Feature (NEW RULES)
_2, 5, 7_
Sequence(s) _____ are missing the <220>Feature and associated headings.
Use of <220> to <223> is MANDATORY if <213>ORGANISM is "Artificial" or "Unknown"
Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 6/01/98, Vol. 63, No. 104, pp. 29631-32)   (Sec. 1.823 of new Rules)

13 _____ PatentIn ver. 2.0 "bug"
Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other means to copy file to floppy disk.

```
RAW SEQUENCE LISTING                      DATE: 01/08/2001
PATENT APPLICATION:  US/09/737,476        TIME: 17:22:26

Input Set : A:\Frenken-seq.txt
Output Set: N:\CRF3\01082001\I737476.raw
```

```
   3 <110> APPLICANT: Unilever plc
   4         Unilever NV
   6 <120> TITLE OF INVENTION: Production of Antibodies
   8 <130> FILE REFERENCE: T7060
C--> 10 <140> CURRENT APPLICATION NUMBER: US/09/737,476
C--> 10 <141> CURRENT FILING DATE: 2000-12-18
  10 <160> NUMBER OF SEQ ID NOS: 67
  12 <170> SOFTWARE: PatentIn version 3.0
  14 <210> SEQ ID NO: 1
  15 <211> LENGTH: 440
  16 <212> TYPE: DNA
  17 <213> ORGANISM: Artificial
  19 <220> FEATURE:
  20 <223> OTHER INFORMATION: VHH with peptide linker
  22 <220> FEATURE:
  23 <221> NAME/KEY: CDS
  24 <222> LOCATION: (1)..(417)
  26 <400> SEQUENCE: 1
  27 cag gtg cag ctg cag gag tca ggg gga gga ttg gtg cag gct ggg ggc       48
  28 Gln Val Gln Leu Gln Glu Ser Gly Gly Gly Leu Val Gln Ala Gly Gly
  29 1               5                   10                  15
  31 tct ctg aga ctc tcc tgt gca gcc tcg gga cgc gcc acc agt ggt cat       96
  32 Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Arg Ala Thr Ser Gly His
  33                 20                  25                  30
  35 ggt cac tat ggt atg ggc tgg ttc cgc cag gtt cca ggg aag gag cgt      144
  36 Gly His Tyr Gly Met Gly Trp Phe Arg Gln Val Pro Gly Lys Glu Arg
  37             35                  40                  45
  39 gag ttt gtc gca gct att agg tgg agt ggt aaa gag aca tgg tat aaa      192
  40 Glu Phe Val Ala Ala Ile Arg Trp Ser Gly Lys Glu Thr Trp Tyr Lys
  41     50                  55                  60
  43 gac tcc gtg aag ggc cga ttc acc atc tcc aga gat aac gcc aag act      240
  44 Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Thr
  45 65                  70                  75                  80
  47 acg gtt tat ctg caa atg aac agc ctg aaa cct gaa gat acg gcc gtt      288
  48 Thr Val Tyr Leu Gln Met Asn Ser Leu Lys Pro Glu Asp Thr Ala Val
  49                 85                  90                  95
  51 tat tat tgt gcc gct cga ccg gtc cgc gtg gat gat att tcc ctg ccg      336
  52 Tyr Tyr Cys Ala Ala Arg Pro Val Arg Val Asp Asp Ile Ser Leu Pro
  53                 100                 105                 110
  55 gtt ggg ttt gac tac tgg ggc cag ggg acc cag gtc acc gtc tcc tca      384
  56 Val Gly Phe Asp Tyr Trp Gly Gln Gly Thr Gln Val Thr Val Ser Ser
  57             115                 120                 125
  59 gaa caa aaa ctc atc tca gaa gag gat ctg aat taataagggc taagctcgaa    437
  60 Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Asn
  61     130                 135
  63 ttc                                                                  440
  66 <210> SEQ ID NO: 2
```

RAW SEQUENCE LISTING   DATE: 01/08/2001
PATENT APPLICATION: US/09/737,476   TIME: 17:22:26

Input Set : A:\Frenken-seq.txt
Output Set: N:\CRF3\01082001\I737476.raw

```
 67 <211> LENGTH: 139
 68 <212> TYPE: PRT
 69 <213> ORGANISM: Artificial ────→     [handwritten: Missing mandatory <220>, <223> features
 71 <400> SEQUENCE: 2                      to explain artificial sequence. See
                                           #12 on Error
                                           Summary Sheet.]
 73 Gln Val Gln Leu Gln Glu Ser Gly Gly Gly Leu Val Gln Ala Gly Gly
 74 1               5                   10                  15
 77 Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Arg Ala Thr Ser Gly His
 78              20                  25                  30
 81 Gly His Tyr Gly Met Gly Trp Phe Arg Gln Val Pro Gly Lys Glu Arg
 82          35                  40                  45
 85 Glu Phe Val Ala Ala Ile Arg Trp Ser Gly Lys Glu Thr Trp Tyr Lys
 86      50                  55                  60
 89 Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Thr
 90 65                  70                  75                  80
 93 Thr Val Tyr Leu Gln Met Asn Ser Leu Lys Pro Glu Asp Thr Ala Val
 94              85                  90                  95
 97 Tyr Tyr Cys Ala Ala Arg Pro Val Arg Val Asp Asp Ile Ser Leu Pro
 98              100                 105                 110
101 Val Gly Phe Asp Tyr Trp Gly Gln Gly Thr Gln Val Thr Val Ser Ser
102          115                 120                 125
105 Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Asn
106      130                 135
109 <210> SEQ ID NO: 3
110 <211> LENGTH: 11
111 <212> TYPE: PRT
112 <213> ORGANISM: Artificial
114 <220> FEATURE:
115 <223> OTHER INFORMATION: myc linker
117 <400> SEQUENCE: 3
119 Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Asn
120 1               5                   10
122 <210> SEQ ID NO: 4
123 <211> LENGTH: 471
124 <212> TYPE: DNA
125 <213> ORGANISM: Artificial
127 <220> FEATURE:
128 <223> OTHER INFORMATION: VHH with linker
130 <220> FEATURE:
131 <221> NAME/KEY: CDS
132 <222> LOCATION: (1)..(459)
134 <400> SEQUENCE: 4
135 cag gtg cag ctg cag cag tca ggg gga ggc ttg gtg cag gct ggg ggg     48
136 Gln Val Gln Leu Gln Gln Ser Gly Gly Gly Leu Val Gln Ala Gly Gly
137 1               5                   10                  15
139 tct ctg aga ctc tcc tgt gta gct tct gaa agc agc ttc agc aac aat     96
140 Ser Leu Arg Leu Ser Cys Val Ala Ser Glu Ser Ser Phe Ser Asn Asn
141              20                  25                  30
143 cac atg ggc tgg tac cgc cgg gct cca ggg aac cag cgc gag ctg gtc    144
144 His Met Gly Trp Tyr Arg Arg Ala Pro Gly Asn Gln Arg Glu Leu Val
```

```
                    RAW SEQUENCE LISTING            DATE: 01/08/2001
                    PATENT APPLICATION:  US/09/737,476    TIME: 17:22:26

                    Input Set : A:\Frenken-seq.txt
                    Output Set: N:\CRF3\01082001\I737476.raw

    145         35                   40                   45
    147 gca act att agt cct ggt ggt agc aca cac tat gta gac tcc gtg aag        192
    148 Ala Thr Ile Ser Pro Gly Gly Ser Thr His Tyr Val Asp Ser Val Lys
    149     50                   55                   60
    151 ggc cga ttc acc atc tcc cga gac aac gcc aag aac aca gtg tat cta        240
    152 Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Asn Thr Val Tyr Leu
    153 65                   70                   75                   80
    155 caa atg gac agc ctg aaa cca gag gac acg gcc gtc tat tac tgt gct        288
    156 Gln Met Asp Ser Leu Lys Pro Glu Asp Thr Ala Val Tyr Tyr Cys Ala
    157                 85                   90                   95
    159 gcc aag ggg agg ggg ctg cag gct atg cag tac tgg ggc cag ggg acc        336
    160 Ala Lys Gly Arg Gly Leu Gln Ala Met Gln Tyr Trp Gly Gln Gly Thr
    161                 100                  105                  110
    163 ctg gtc acc gtc tcc tca gcg cac cac agc gaa gac ccc agc tcc gcg        384
    164 Leu Val Thr Val Ser Ser Ala His His Ser Glu Asp Pro Ser Ser Ala
    165             115                  120                  125
    167 gcc gcc cat cac cat cac cat cac ggg gcc gca gaa caa aaa ctc atc        432
    168 Ala Ala His His His His His His Gly Ala Ala Glu Gln Lys Leu Ile
    169     130                  135                  140
    171 tca gaa gag gat ctg aat ggg gcc gca tagtaacaat tg                      471
W-->172 Ser Glu Glu Asp Leu Asn Gly Ala Ala
    173 145                  150
    176 <210> SEQ ID NO: 5
    177 <211> LENGTH: 153
    178 <212> TYPE: PRT
    179 <213> ORGANISM: Artificial
    181 <400> SEQUENCE: 5
    183 Gln Val Gln Leu Gln Gln Ser Gly Gly Gly Leu Val Gln Ala Gly Gly
    184 1               5                   10                  15
    187 Ser Leu Arg Leu Ser Cys Val Ala Ser Glu Ser Ser Phe Ser Asn Asn
    188             20                  25                  30
    191 His Met Gly Trp Tyr Arg Arg Ala Pro Gly Asn Gln Arg Glu Leu Val
    192             35                  40                  45
    195 Ala Thr Ile Ser Pro Gly Gly Ser Thr His Tyr Val Asp Ser Val Lys
    196     50                  55                  60
    199 Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Asn Thr Val Tyr Leu
    200 65                  70                  75                  80
    203 Gln Met Asp Ser Leu Lys Pro Glu Asp Thr Ala Val Tyr Tyr Cys Ala
    204             85                  90                  95
    207 Ala Lys Gly Arg Gly Leu Gln Ala Met Gln Tyr Trp Gly Gln Gly Thr
    208             100                 105                 110
    211 Leu Val Thr Val Ser Ser Ala His His Ser Glu Asp Pro Ser Ser Ala
    212             115                 120                 125
    215 Ala Ala His His His His His His Gly Ala Ala Glu Gln Lys Leu Ile
    216     130                 135                 140
    219 Ser Glu Glu Asp Leu Asn Gly Ala Ala
    220 145                 150
    223 <210> SEQ ID NO: 6
    224 <211> LENGTH: 468
```

Missing <220>, <223>
See #12

RAW SEQUENCE LISTING   DATE: 01/08/2001
PATENT APPLICATION: US/09/737,476   TIME: 17:22:27

Input Set : A:\Frenken-seq.txt
Output Set: N:\CRF3\01082001\I737476.raw

```
225 <212> TYPE: DNA
226 <213> ORGANISM: Artificial
228 <220> FEATURE:
229 <223> OTHER INFORMATION: VHH with linker
231 <220> FEATURE:
232 <221> NAME/KEY: CDS
233 <222> LOCATION: (1)..(456)
235 <400> SEQUENCE: 6
236 cag gtg cag ctg cag gag tct ggg gga ggc ctg gtg cag gct ggg ggg      48
237 Gln Val Gln Leu Gln Glu Ser Gly Gly Gly Leu Val Gln Ala Gly Gly
238 1               5                  10                  15
240 tct ctg aga ctc tcc tgt gta gcc tct gga aac acc ttc agt atc ata      96
241 Ser Leu Arg Leu Ser Cys Val Ala Ser Gly Asn Thr Phe Ser Ile Ile
242             20                  25                  30
244 gct atg gcc tgg tac cgc cag gct cca ggg aag cag cgc gag gtg gtc     144
245 Ala Met Ala Trp Tyr Arg Gln Ala Pro Gly Lys Gln Arg Glu Val Val
246         35                  40                  45
248 gca agt att aat agt att ggc agc aca aat tat gca gac tcc gtg aag     192
249 Ala Ser Ile Asn Ser Ile Gly Ser Thr Asn Tyr Ala Asp Ser Val Lys
250     50                  55                  60
252 ggg cga ttc acc atc tcc aga gac aac gcc aag aac aca gtg tat ctg     240
253 Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Asn Thr Val Tyr Leu
254 65                  70                  75                  80
256 caa atg agc agc ctg aaa cct gag gac acg gcc gtc tat tac tgt gct     288
257 Gln Met Ser Ser Leu Lys Pro Glu Asp Thr Ala Val Tyr Tyr Cys Ala
258             85                  90                  95
260 gcc ggt aat ttg ctg gtt aag agg cct tac tgg ggc cag ggg acc ctg     336
261 Ala Gly Asn Leu Leu Val Lys Arg Pro Tyr Trp Gly Gln Gly Thr Leu
262             100                 105                 110
264 gtc acc gtc tcc tca gaa ccc aag aca cca aaa cca caa cca gcg gcc     384
265 Val Thr Val Ser Ser Glu Pro Lys Thr Pro Lys Pro Gln Pro Ala Ala
266         115                 120                 125
268 gcc cat cac cat cac cat cac ggg gcc gca gaa caa aaa ctc atc tca     432
269 Ala His His His His His His Gly Ala Ala Glu Gln Lys Leu Ile Ser
270     130                 135                 140
272 gaa gag gat ctg aat ggg gcc gca tagtaacaat tg                       468
W--> 273 Glu Glu Asp Leu Asn Gly Ala Ala
274 145                 150
277 <210> SEQ ID NO: 7
278 <211> LENGTH: 152
279 <212> TYPE: PRT
280 <213> ORGANISM: Artificial  ———>
282 <400> SEQUENCE: 7
284 Gln Val Gln Leu Gln Glu Ser Gly Gly Gly Leu Val Gln Ala Gly Gly
285 1               5                  10                  15
288 Ser Leu Arg Leu Ser Cys Val Ala Ser Gly Asn Thr Phe Ser Ile Ile
289             20                  25                  30
292 Ala Met Ala Trp Tyr Arg Gln Ala Pro Gly Lys Gln Arg Glu Val Val
293         35                  40                  45
```

Missing <220>, <223>
See #12

RAW SEQUENCE LISTING   DATE: 01/08/2001
PATENT APPLICATION: US/09/737,476   TIME: 17:22:27

Input Set : A:\Frenken-seq.txt
Output Set: N:\CRF3\01082001\I737476.raw

```
296 Ala Ser Ile Asn Ser Ile Gly Ser Thr Asn Tyr Ala Asp Ser Val Lys
297     50                  55                  60
300 Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Asn Thr Val Tyr Leu
301 65                  70                  75                  80
304 Gln Met Ser Ser Leu Lys Pro Glu Asp Thr Ala Val Tyr Tyr Cys Ala
305                 85                  90                  95
308 Ala Gly Asn Leu Leu Val Lys Arg Pro Tyr Trp Gly Gln Gly Thr Leu
309                 100                 105                 110
312 Val Thr Val Ser Ser Glu Pro Lys Thr Pro Lys Pro Gln Pro Ala Ala
313     .   115                 120                 125
316 Ala His His His His His Gly Ala Ala Glu Gln Lys Leu Ile Ser
317     130                 135                 140
320 Glu Glu Asp Leu Asn Gly Ala Ala
321 145                 150      .
324 <210> SEQ ID NO: 8
325 <211> LENGTH: 462
326 <212> TYPE: DNA
327 <213> ORGANISM: Artificial
329 <220> FEATURE:
330 <223> OTHER INFORMATION: VHH with linker
332 <220> FEATURE:
333 <221> NAME/KEY: CDS
334 <222> LOCATION: (1)..(450)
336 <400> SEQUENCE: 8
337 acc atg gcc cag gtg aaa ctg cag cag tct ggg gga gga ttg gtg cag      48
338 Thr Met Ala Gln Val Lys Leu Gln Gln Ser Gly Gly Gly Leu Val Gln
339 1               5                   10                  15
341 gct ggg ggc cct ctg agg ctc tcc tgt gca gcc tct gga cgc acc ttc      96
342 Ala Gly Gly Pro Leu Arg Leu Ser Cys Ala Ala Ser Gly Arg Thr Phe
343                 20                  25                  30
345 agt aac tat gcc gtg ggc tgg ttc cgc cag gct cca ggg aag gag cgt     144
346 Ser Asn Tyr Ala Val Gly Trp Phe Arg Gln Ala Pro Gly Lys Glu Arg
347             35                  40                  45
349 gag ttt gtc gct gct att agc cgt gat ggt ggg cgc aca tac tat gcg     192
350 Glu Phe Val Ala Ala Ile Ser Arg Asp Gly Gly Arg Thr Tyr Tyr Ala
351         50                  55                  60
353 gac tcc gtg aag ggc cga ttc gcc gtc tcc aga gac tac gcc gag aac     240 .
354 Asp Ser Val Lys Gly Arg Phe Ala Val Ser Arg Asp Tyr Ala Glu Asn
355 65                  70                  75                  80
357 acg gtg tat ctg caa atg aac agc ctg aaa cct gag gac acg gcc gtt     288
358 Thr Val Tyr Leu Gln Met Asn Ser Leu Lys Pro Glu Asp Thr Ala Val
359                 85                  90                  95
361 tat tac tgt aac aca agg gcc tac tgg ggc cag ggg acc cag gtc acc     336
362 Tyr Tyr Cys Asn Thr Arg Ala Tyr Trp Gly Gln Gly Thr Gln Val Thr
363                 100                 105                 110
365 gtc tcc tca gcg cac cac agc gaa gac ccc agc tcc gcg gcc gcc cat     384
366 Val Ser Ser Ala His His Ser Glu Asp Pro Ser Ser Ala Ala Ala His
367         115                 120                 125
369 cac cat cac cat cac ggg gcc gca gaa caa aaa ctc atc tca gaa gag     432
```

VERIFICATION SUMMARY   DATE: 01/08/2001
PATENT APPLICATION: US/09/737,476   TIME: 17:22:28

Input Set : A:\Frenken-seq.txt
Output Set: N:\CRF3\01082001\I737476.raw

L:10 M:270 C: Current Application Number differs, Replaced Current Application No
L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:172 M:361 W: Invalid Split Codon, Sequence data for SEQ ID#: 4
L:273 M:361 W: Invalid Split Codon, Sequence data for SEQ ID#: 6
L:374 M:361 W: Invalid Split Codon, Sequence data for SEQ ID#: 8
L:475 M:361 W: Invalid Split Codon, Sequence data for SEQ ID#: 10
L:1684 M:361 W: Invalid Split Codon, Sequence data for SEQ ID#: 66